

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 Computer Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2003, 10:46:02 : Search time 2998 Seconds
(without alignments)
11409.229 Million cell updates/sec

Title: us-09-635-949-33

Period score: 2112

Sequence: 1 aaaaaggaatccaggaaccc.....ttcttctggcctgtttgt 2112

Scoring table: IDENTITY_NUC

Gapop 10_0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32408132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: EST:

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estic:*

9: qb_est1:*

10: qb_est2:*

11: qb_est3:*

12: qb_est4:*

13: qb_est5:*

14: qb_est6:*

15: em_estatun:*

16: em_estom:*

17: qb_qss:*

18: em_qss_hum:*

19: em_qss_inv:*

20: em_qss_pln:*

21: em_qss_vit:*

22: em_qss_fum:*

23: em_qss_wan:*

24: em_qss_mus:*

25: em_qss_other:*

26: em_qss_pro:*

27: em_qss_red:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	532.8	25.2	641	13	RG986976
2	531.4	25.2	602	10	AW268636
3	516.8	24.5	768	13	BI652094
4	502	23.8	935	14	RG884973
5	464	22.0	949	13	BI108493
6	444.2	21.0	457	10	AW607534

c	7	449.2	20.8	457	12	BF924943
	8	437.6	20.7	861	13	BF557089
	9	431.6	20.4	499	9	AI1840826
c	10	424	20.1	583	14	BQ358250
c	11	416.4	19.7	461	10	BE175841
	12	409.8	19.4	541	10	AW762810
	13	404	19.1	636	12	BF308711
	14	401.4	19.0	493	12	BF654270
	15	377	17.9	721	13	BI250805
	16	371.4	17.6	445	10	AW762749
c	17	364.4	17.3	405	14	BQ370594
	18	353.2	16.7	420	9	AA913032
c	19	345.8	16.4	508	12	BF329641
	20	329.8	15.6	543	9	AI527763
	21	327.8	15.5	455	10	AW346379
c	22	327.6	15.5	433	13	BI005418
c	23	319.6	15.1	473	14	BI005408
	24	312.8	14.8	510	9	AI786568
	25	312.6	14.8	550	9	AI316361
c	26	311.8	14.8	677	10	BE654123
c	27	310	14.7	400	10	AW580493
c	28	297.6	14.1	685	9	AI907471
c	29	296.4	14.0	438	13	RG996753
c	30	293.6	13.9	466	17	AZ814570
c	31	286.8	13.6	399	10	AW580482
	32	282.6	13.4	450	13	RG995163
c	33	282.4	13.4	325	12	BI012691
	34	280	13.3	844	12	BF122029
	35	278.8	13.2	415	14	RG996752
	36	277	13.1	1006	9	AI554449
	37	277	13.1	1011	9	AI544560
	38	277	13.1	1046	9	AI550203
	39	276.2	13.1	876	9	AI552647
	40	276.2	13.1	934	9	AI544630
	41	274.2	13.0	896	14	BQ736956
	42	274	13.0	1114	9	AI545700
	43	272.6	12.9	1007	9	AI547279
	44	266.6	12.6	994	9	AI514052
	45	260.4	12.3	594	17	AQ427707

ALIGNMENTS

RESULT 1	RG986976	641 bp	mRNA	linear	EST 18 JUN 2003
LOCUS	RG986976	MG:HT1183-050101-003-d05	HT1183	Homo sapiens cDNA, mRNA sequence	
DEFINITION	MG:HT1183-050101-003-d05				
ACCESSION	RG986976				
VERSION	RG986976.1				
KEYWORDS	EST				
SOURCE	human				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 641)				
AUTHORS	Dias Neto F., Garcia Corredor K., Verjovski Almeida S., Brites M., Dias Neto F., da Silva W. Jr., Zago M.A., Bordin S., Costa F., Goldman G.H., Carvalho A.F., Matsukuma A., Jbara G.S., Simpson B.H., Branstetter A., de Oliveira P.S., Buchet P., Junqueira C.V., O'Hare M.J., Soares F., Brentani R., Reis L.F., de Souza S.J. and Simpson A.J.				
TITLE	Shetun sequence of the human transcriptome with ORF expressed sequence tags				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.				
MEDLINE	2020263				
COMMENT	contact: Simpson A.J.3. Laboratory of Cancer Genetics Indiana Institute for Cancer Research Box Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11 2704922 Fax: +55-11 2707001				

```

RESOLUTION 2
AWZ6R6.66
LORDS
DEPTIN TON
AWZ6R6.66 602 UP mRNA 11091 EST 05 JAN 2000
XVAL11.X1 SQUASH-NET-2 GAC ST HOMO SAPIENS CHINA CHINA
IMACT:481474 47811141 TO SW_FILZ MODULE 1-17869 FILLIN 2
PREPENSE-OR, 3, mRNA squashed,
AWZ6R6.66
AWZ6R6.66
AWZ6R6.66 611665666
KEYWORDS
EST

```

[illegible]

FIGURE 5
 LOCALIZATION/QUALITY
 OF FINGER-4000 FROM TAPERS
 HIGH QUALITY SPECTROSCOPY: 4000
 OF FINGER-4000 FROM TAPERS
 POSSIBLE TO OBTAIN INFORMATION
 IMAGE CONSUMPTION (INFO-IMAGE) FOR FURTHER INFORMATION
 THIS VALUE IS AVAILABLE TOGETHER WITH THROUGH LINE CONTACT THE

[illegible][illegible][illegible][illegible][illegible]

.....

146 250 AAGTGTATCTGCTCAACGGATATATGCTCAAGCTGATGGTCTCTCTTCAAGTGGCTTGA 309
 QY 648 CTTCTTCATGCAAAAGCTGTCAGTATGGTGTGATGTTGTATAAGAGCAAAATACGCTGAC 707
 Db 310 CTTCTTCATGCAAAAGCTGTCAGTATGGTGTGATGTTGTATAAGAGCAAAATACGCTGAC 369
 QY 708 AATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 767
 Db 370 AATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429
 QY 768 AATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 827
 Db 430 AATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 489
 QY 828 GTTACATCTGCTAAGTGTGCTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887
 Db 490 GTTACATCTGCTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
 QY 888 GTTACATCTGCTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
 Db 550 GTTACATCTGCTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602

RESULT 3
 R1652094
 LOCUS 603298884F1 NC1_GAP_Man3 768 bp mRNA linear EST 12-SEP-2001
 DEFINITION mRNA sequence.

ACCESSION R1652094
 VERSION R1652094.1 GI:15556430
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 768)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: capbs-r@mail.nih.gov
 Tissue Procurement: Lotmar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Inceyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/[LLNL](http://image.llnl.gov) at:
<http://image.llnl.gov>
 Plate: LLNL1862 row: 4 column: 21
 High quality sequence stop: 767.

FEATURES
 Source
 Location/Qualifiers
 1..768
 organism "Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone "IMAGE:5335564"
 /clone_lib="NCI_GAP_Man3"
 /tissue_type="tumor, gross tissue"
 /dev_stage "10 months"
 /lab_host="DH10B"
 /note "organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: of 140 bp.
 Library constructed by Life Technologies. Investigators
 provided samples: Lotmar Hennighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."

BASE COUNT 204 a 204 c 212 g 148 t
 ORIGIN

Query Match

Best local Similarity 86.6% Ident. No. 140-122;

Matches 627; conservative 0; Mismatches 92; Indels 5; Gaps 5;

QY 1257 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1314
 Db 1 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 63
 QY 1417 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1474
 Db 63 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 120
 QY 1477 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1434
 Db 121 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 140
 QY 1436 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1493
 Db 181 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 240
 QY 1496 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1553
 Db 241 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 300
 QY 1556 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1614
 Db 301 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 360
 QY 1615 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1674
 Db 361 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 420
 QY 1675 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1734
 Db 421 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 480
 QY 1735 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1794
 Db 481 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 540
 QY 1794 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1853
 Db 541 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 600
 QY 1854 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1914
 Db 601 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 660
 QY 1914 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 1971
 Db 661 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT 720
 QY 1972 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT
 Db 721 TAACTAGAGCTTAAAGATATCTCTATATCTATATCTATATCTATATCTATATCTATATCT

R1652094
 LOCUS 603298884F1 NC1_GAP_Man3 768 bp mRNA linear EST 12-SEP-2001
 DEFINITION mRNA sequence.

ACCESSION R1652094
 VERSION R1652094.1 GI:15556430
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 935)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: capbs-r@mail.nih.gov
 Tissue Procurement: Mark Macdonald, Ph.D. and Nancy L. Florenzo, Ph.D.

